



LARGE SCALE, DATA INTENSIVE COMPUTING

SC 98 Tutorial

Introduction

BILL KRAMER

DEPUTY DIVISION DIRECTOR

DEPARTMENT HEAD, HIGH PERFORMANCE COMPUTING DEPARTMENT

kramer@nersc.gov

510-486-7577

Agenda

- **Introduction : Bill Kramer**
- **Motivation with real world example (STAR) : Craig Tull**
- **Optimizing access to scientific data : Arie Shoshani**
- **Mass storage (HPSS) : Harvard Holmes**
- **Distributed storage (DPSS) : Brian Tierney**
- **Clipper: Bob Lucas**
- **General Question Session: All**

What is Data Intensive Computing?

- **Multiple styles of computing have evolved**
 - **Computationally Intensive**
 - **Data Analysis**
 - **Data Base**
 - **Distributed Computing**
 - **Data Presentation**

Computationally Intensive

- **Simulation of physical properties**
- **Focus on Floating Point calculations**
- **Can be highly parallel**
- **Well defined mathematics**
- **Can produce large amounts of data**
 - typically large files with internal data structures
 - data can be recreated
- **Usually not real time**
- **Often coupled with post processing visualization**
- **Examples**
 - CFD Climate
 - QCD Structural Analysis

Data Analysis

- Goal is to understand, catalogue and evaluate existing data
- Data may be newly acquired
 - Experimental
 - Observational
- Often discrete files
 - Moderate size of files
 - Large number of files
- Integer and floating point
- Examples
 - Satellite data
 - Weather Observations
 - Accelerator data
 - Medical Images

- Goal is data storage, relationships and retrieval
- Little use to date in scientific contexts
 - Mostly business driven
- Focus on sorting and manipulation, not floating point
- Record oriented rather than file oriented
 - Often in a structured environment
- Examples
 - Transaction processing
 - Banking
 - Personnel

Distributed Computing

- Having data and/or computation at multiple physical locations
- Sharing resources
- Relies on commodity infrastructure
- Cooperative processing
- Examples
 - PVM/MPI
 - AFS/DFS/NFS
 - “Client Server”

Data Presentation

- **Goal is to present data in an understandable manner**
 - Turn data into information
- **Visualization**
 - Gaining insight from the combination and manipulation of data
 - Time accurate visualization
 - Real time
- **Web browsing**

Data Intensive Computing

- The general definition is the merging of all these styles
 - Degrees vary
- Attributes
 - Performance limitation are more I/O and computation
 - ✱ Ratio of computation to I/O
 - Arrangement of data with high floating point computation
 - Data not necessarily co-located with computation
 - Simulation and data analysis integrated
 - Presentation of data integrated with creation and analysis

- **Several Experiments on line in 1999**
 - **SLAC B Factory – 100 TB Data/yr**
 - **D0 and CDF at Fermilab – 100 TB/yr**
 - **RHIC 1,500 TB/yr (only capture 50% of raw data)**
- **Data reuse ranges from 10% HEP to 40%**
- **Experiments run 5–10 years**
- **Cost is the only factor constraining storage**
 - **Processing power is secondary**
- **OODB important, relational DB not workable**
- **Total budget in US = \$1B/yr**

- Example for one CFD group – about 1.9% of the NERSC system.
- 25 MB/Gflop for this application
- Total requirement for Storage = 27 TB/yr
- Computational increase of NERSC system to 100 – 1,000 GFlops in 10 years.
- Storage would need to be 3 to 30 PB year

Humane Genome

- **5 year projection 100 TB of data/site – with 10 major sites for the Human Genome**
- **Raw data discarded – 1,000 times more than stored**
- **Expansion**
 - **More species (Mouse, Fly, Yeast, Zebra Fish, etc.)**
 - **Individual**
 - × **1 Billion people, 1 % of their genomes/yr**
 - × **1 PB/yr**

- **Crystallography Today**
 - 10–20 major facilities in US
 - 2–4 beam lines per facility
 - 10 MB/s data production
 - 100 TB/yr
 - many small systems
- **Protein Engineering**
- **Seismology**
- **Cosmology**
- **Medical Records/Imaging**

- **Data organization for efficient retrieval of Large Datasets**
 - Parallel Disks
 - Tape organization
- **Data Management Tools for scientific applications**
- **Data Structures and operations for scientific applications**
- **Application Areas to date**
 - Humane Genome
 - Climate Modeling
 - Terrain Navigation
 - Geophysical Wells/Boreholes
 - High Energy/Nuclear Physics
- **Distributed Parallel Storage System (DPSS)**

Analyzing Climate Variability with Ensembles of Simulations

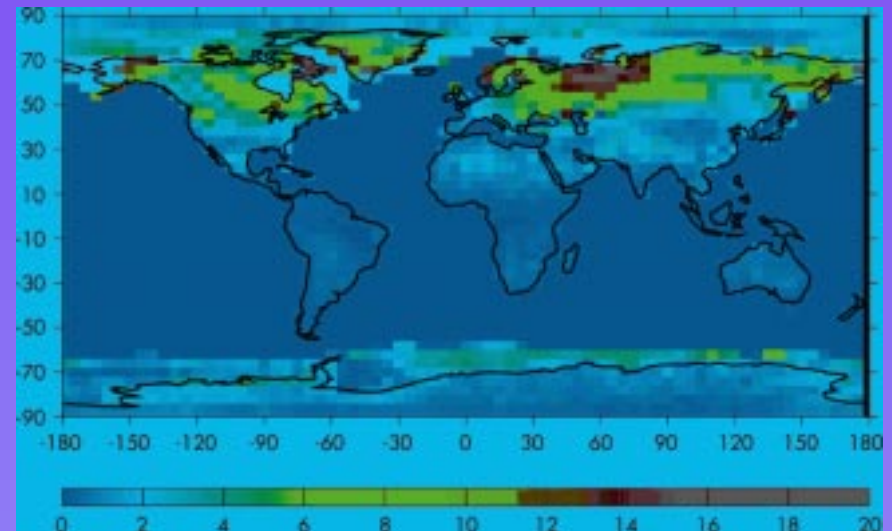
Gerald Potter and Michael Wehner, Lawrence Livermore National Laboratory

Research Objectives

Climate—the average of weather—can be simulated by performing long-term integrations of weather and averaging the results. The details of the weather are not actually realized; however, given a realistic-enough model, we may generate climate statistics that can be legitimately compared with observations of the true climate system.

Significance

Long-term climate simulations tax the capabilities of even the most powerful computers. It is desirable to increase the resolution of global models substantially in order to better simulate regional features. Comprehensive climate models also must include processes other than atmospheric circulation, such as ocean circulation, sea ice processes, biological processes, and atmospheric chemistry.



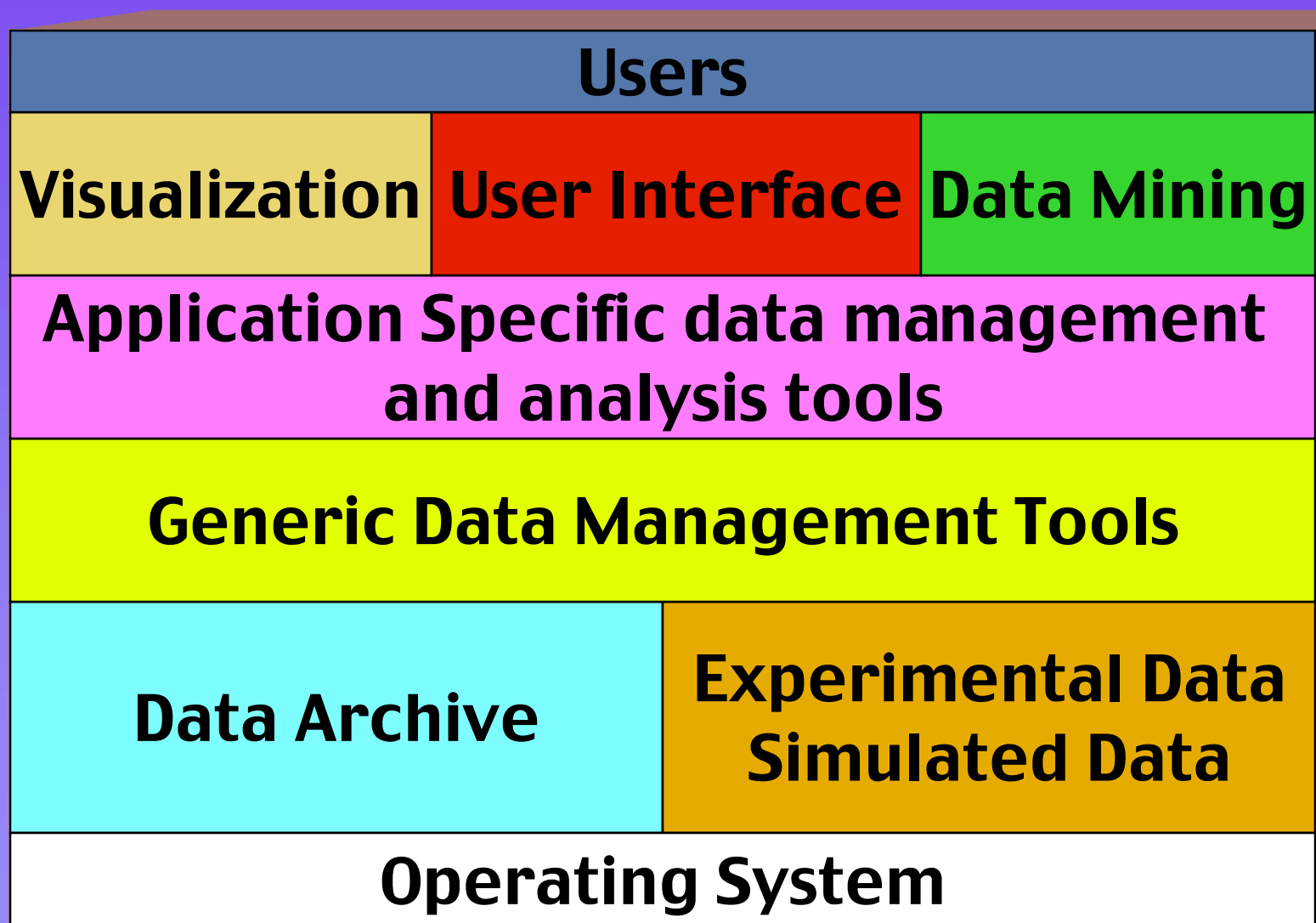
The number of atmospheric general circulation model calculations required to compute the decadal averaged seasonal surface temperature to within 0.5° Kelvin at 95% statistical certainty.

Challenges to Data Intensive Computing

- **System Balance**
 - **I/O vs Computation**
 - × **Within a computer**
 - × **Between a system of computers**
 - × **Between sites**

- **Example – HiPPI**
 - **At NERSC**
 - ✧ 14 HiPPI connections from compute systems
 - ✧ 8 HiPPI connections to Max Strat Raid Disk
 - ✧ 14 HiPPI connections to HPSS Servers
 - **Performance per link**
 - ✧ Peak 100 MB/s
 - ✧ IPI-3 Third Party from C-90 to Disk 60 MB/s
 - ✧ TCP/IP – Cray Memory 50–60 MB/s
 - ✧ Compute to storage servers (mem) 14 MB/s
 - ✧ Compute to storage disk 4–5 MB/s
 - ✧ User experience 1–2 MB/s
- **DPSS – 10 MB/s per server (for N servers)**

How to think of Data Intensive Computing



What is critical

- **Data Intensive Computing requirements will influence the eventual solutions of the other areas**
- **Clusters of SMPS**
 - **Internal I/O bandwidth**
 - **Non-uniform access costs**
 - **Bottlenecks**
- **Computational Grids**
 - **Data Locality**
 - **Use of Databases rather than files structures**
- **Data bases integrated with hierarchical archive systems**

Summary



- **Data is of increasing importance**
- **Cost of Data Storage is a constraint in many areas**
- **Usage of traditional technologies is not sufficient to accomplish scientific mission**
- **Investment in data intensive computing research and development is imperative**